Q2. Discuss solution of ill-posed systems of equations with regard to methods

Well (ill)-Posed Problems

- The quality of solution depends on
  - the problem itself and
  - the computer
- According to Hadamard, a problem is well-posed if the solution
  - exists
  - is unique and
  - depends continuously on the data (stable).
- Typically practical inverse problems are all ill-posed.
- Even well-posed problems may be unstable or ill-conditioned when implemented in digital computers.

Solutions for Linear Systems

- A linear system of equations in a matrix form:

$$Ax = b$$

where the coefficient matrix $A \in \mathbb{R}^{m \times n}$, the constant vector $b \in \mathbb{R}^m$ and the variable vector $x \in \mathbb{R}^n$.

- Suppose $m = n$. The solution

$$\hat{x} = A^{-1}b$$

may be disastrous especially when n is large.

- Types of the solution:
  - No solution
  - Multiple solutions
- Solving methods: decompositions and regularization

Cholesky Decomposition

- Suppose the matrix A is
  - Symmetric and
  - Positive definite
- Decompose A into a unique lower and upper triangular matrices:

$$A = LL^T$$

- On the LHS, we have

$$A\hat{x} = LL^T\hat{x} = L(L^T)\hat{x}$$

- Solve by the forward and backward substitutions:

$$\begin{aligned} Ly &= b \\ L^T\hat{x} &= y \end{aligned}$$

- fast, stable and requires less space!

Truncated SVD

- Decomposes any matrix A as

$$A \;=\; UDV^T,$$

where U and V are orthogonal matrices such that $U^TU = V^TV = I$; D is diagonal with singular values of A.

- If A is non-singular, we write

$$A^{-1} \;=\; V\Sigma^{-1}U^T,$$

where $\Sigma^{-1} = [\text{diag}(\frac{1}{\sigma_i})]$.

Overdetermined Systems

- More equations than unknowns
- b does not lie in $\mathscr{R}(A) \Rightarrow$ No solution
- Yields the unique solution by minimizing the residual $||Ax - b||_2$.
- Using the SVD, we write

$$\begin{aligned} \min ||Ax - b|| &= \min ||U\Sigma V^T - b|| \\ &= \min ||\Sigma V^T x - U^T b|| \\ &= \min ||\Sigma v - \tilde{b}||, \end{aligned}$$

where $v = V^T x$ and $\tilde{b} = U^T b$.

- The min. length solution for v is

$$v = \Sigma^+ \tilde{b}.$$

- Hence

$$\hat{x} = V\Sigma^+\tilde{b} = V\Sigma^+U^Tb.$$

- Summary:
  - Compute the SVD of A : $A = U\Sigma V^T$
  - Zero-out 'small' σ_i's of Σ.
  - Obtain

$$\hat{x} = V\Sigma^+(U^Tb),$$

where $\Sigma^+ = [\text{diag}(\frac{1}{\sigma_i})]$.

Underdetermined Systems

- Effectively, fewer equations than unknowns
- $b \in \mathscr{R}(A) \Rightarrow$ Multiple solutions
- We may choose the smallest norm solution similarly to the overdetermined case.
- Pros:
  - Robust when A is singular or near singular
  - Treats both the underdetermined and overdetermined systems identically
- Cons:
  - Computational more demanding.

Regularized LS method

- To improves the stability, add regularization in the minimization:

$$\hat{x} = \arg\min ||Ax - b||^2 + ||\Gamma x||^2$$

  where Γ is the regularization matrix or *Tikhonov* matrix

- Regularized solution:

$$\hat{x} \;=\; (A^T A + \Gamma^T \Gamma)^{-1} A^T b$$

- $\Gamma = 0 \Rightarrow$ Conventional LS solution.

Choice of Γ Using the KF Theory

- Perceive the following to be the measurement equation:

$$b_k = Ax_k + w_k$$

- Suppose $\hat{x}_{k|k-1} \sim \mathcal{N}(0, \sigma_x^2 I)$, and $w_k \sim \mathcal{N}(0, \sigma_b^2 I)$.
- Then the updated state:

$$\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + W(b_k - \hat{b}_{k|k-1}) \\
\hat{x} &= Wb \\
&= \left[A^T(\sigma_b^2 I)^{-1}A + (\sigma_x^2 I)^{-1}\right]^{-1} A^T (\sigma_b^2 I)^{-1} b \\
&= \frac{1}{\sigma_b^2}\left[\frac{1}{\sigma_b^2} A^T A + \frac{1}{\sigma_x^2} I\right]^{-1} A^T b \\
&= \left[A^T A + \left(\frac{\sigma_b}{\sigma_x}\right)^2 I\right]^{-1} A^T b
\end{aligned}$$

- The above expression suggests to choose Γ to be $\Gamma = \alpha I$, where $\alpha = \frac{\sigma_b}{\sigma_x}$.

Pseudo-inverses

- works well for full-rank matrix A.
- Case (i): Overdetermined systems
  - Yields the unique solution in the minimum residual, $||Ax - b||$ sense.
  - Write

$$A^T A\hat{x} \;=\; A^T b$$

  - As $A^T A$ is non-singular, we get

$$\hat{x} \;=\; A^+ b,$$

where the pseudo-inverse matrix

$$A^+ \;=\; (A^T A)^{-1} A^T.$$

Pseudo-inverse (Cont'd)

- Case (ii): Underdetermined systems
  - Yields the unique solution in the smallest length, $||x||$ sense:

$$\hat{x} = A^{+}b,$$

where the pseudo-inverse matrix

$$A^{+} = A^{T}(AA^{T})^{-1}.$$

- Limitations:
  - $A^{T}A$ may be singular or near-singular
  - matrix-squared form may amplify roundoff errors !
  - Remedy: Use the SVD on $A^{T}A$ or the QR on A directly.

QR Decomposition in Pseudo-inverses

- Decompose A into

$$A = QR,$$

where R is upper triangular; Q is orthogonal such that $QQ^T = I$.

- For an overdetermined case,

$$\begin{aligned} \hat{x} &= (A^T A)^{-1} A^T b \\ &= (R^T Q^T Q R)^{-1} R^T Q^T b \\ &= (R^T R)^{-1} R^T Q^T b = R^{-1} Q^T b \\ \Rightarrow R\hat{x} &= \underbrace{Q^T b}_{\text{rotate}} \end{aligned}$$

- Use back substitution to get the stable solution.

References

- J. Reilly, *ECE 712:Matrix Computations for Signal Processing*, Course notes.
- G. Golub and C. Van Loan, *Matrix Computations*, John Hopkins, 1996.
- T. Moon and W. Stirling, *Mathematical Methods and Algorithms*, Prentice-Hall, 2001.

Thank you!